14th International Conference on Information Fusion
Chicago, Illinois, USA, July 5-8, 2011

Multi-Sensor Data Fusion: An Unscented Least Squares Approach

Jemin George and Lance M. Kaplan
{jemin.george & lance.m.kaplan}@us.army.mil
U.S. Army Research Laboratory, Adelphi, MD 20783-1197

***Abstract*—This manuscript provides an approach to solve the nonlinear least squares problem that arises in decentralized fusion. Even though almost all sensor noise can be modeled as additive noise, the additive nature of the measurement noise is lost when the signal is processed at the sensor node. The proposed approach employs the unscented transformation before the estimation problem at the central node is posed as a nonlinear least squares problem. Numerical simulations indicate that the proposed unscented transformation based approach yields desired results.**

Keywords: Sensor fusion, decentralized fusion, unscented transformation, nonlinear least squares, iterative least squares

I. Introduction

This manuscript provides an unscented approach to the nonlinear least squares problem that arises in decentralized fusion [1] where only the processed signal is sent to the central node as oppose to the raw measurements. Due to the signal processing conducted at the sensor node, one might loose the additive nature of the measurement noise. For example, consider the i^{th} sensor observations of the following form

$$\tilde{\mathbf{y}}_i = \mathbf{f}(\tilde{\mathbf{x}}) + \mathbf{v}_i \tag{1}$$

where $\tilde{\mathbf{y}}_i \in \Re^m$ are the measurements, $\tilde{\mathbf{x}} \in \Re^n$ are the ground truth quantities (or parameters) of interest, and $\mathbf{v}_i \in \Re^m$ indicates the additive measurement noise. In a decentralized approach, the measurements obtained at the sensor node are first processed before they are relayed to a central processing node. Let $\tilde{\mathbf{z}}_i \in \Re^r$ be the processed signal that is sent to the central node from the i^{th} sensor, i.e.,

$$\tilde{\mathbf{z}}_i = \mathbf{h}_i\left(\tilde{\mathbf{y}}_i\right) = \mathbf{h}_i\left(\mathbf{f}(\tilde{\mathbf{x}}) + \mathbf{v}_i\right) + \mathbf{w}_i \tag{2}$$

where $\mathbf{w}_i \in \Re^r$ indicates some additive noise. In general, the nature of the nonlinear function $\mathbf{h}_i(\cdot)$ can change for each sensor. At the central node, the measurements from N sensor nodes are fused to yield a superior estimate of the quantity of interest. In least squares approach, the estimates are obtained as the hypothesized parameters that minimize the least squares cost function,

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} J(\mathbf{x}) \tag{3}$$

where

$$J(\mathbf{x}) = \frac{1}{2N}\sum_{i=1}^{N}\left\{\tilde{\mathbf{z}}_i - \hat{\mathbf{z}}_i(\mathbf{x})\right\}^T\left\{\tilde{\mathbf{z}}_i - \hat{\mathbf{z}}_i(\mathbf{x})\right\} \tag{4}$$

where $\hat{\mathbf{z}}_i(\mathbf{x})$ indicates the estimated processed signal for the hypothesized parameter $\mathbf{x}$. Even though the estimation criteria given in (4) can be modified such that the residuals are weighted to yield a weighted least squares approach [2], here we only consider the simple least squares. For traditional least squares, the measurement is such that the additive noise term appears outside the nonlinear function, i.e., $\tilde{\mathbf{z}}_i = \mathbf{h}_i(\mathbf{f}(\tilde{\mathbf{x}})) + \mathbf{w}_i$. Then the best choice for the estimated processed signal for hypothesized $\mathbf{x}$ is

$$\hat{\mathbf{z}}_i(\mathbf{x}) = \mathbf{h}_i\left(\mathbf{f}(\mathbf{x})\right) \tag{5}$$

Now, the minimization of (4) is equivalent to the calculation of the maximum likelihood (ML) estimate of $\mathbf{x}$, given that (4) is a weighted least squares cost function. For general nonlinear mappings $\mathbf{h}_i(\cdot)$, there exists no systematic approach to select $\hat{\mathbf{z}}_i(\mathbf{x})$. Once a proper expression for the predicted processed measurement is selected for the hypothesized $\mathbf{x}$, one can choose several traditional approaches to minimize (4). Two examples of these traditional approaches include the zero-order function minimization schemes, such as the simplex algorithm, and the first-order steepest decent approaches, such as the iterative linear least squares. As presented in this paper, the performance of these approaches depends on the mappings $\mathbf{h}_i(\cdot)$ and the predicted processed measurement expression, $\hat{\mathbf{z}}_i(\mathbf{x})$, selected.

This paper presents an unscented transformation [3]–[5] based approach to the nonlinear least squares problem that arises in decentralized data fusion. The proposed approach can be used to improve the solution to the nonlinear least squares problem in (4). Proposed unscented transformation based approach utilizes the *sigma points* to represent $\hat{\mathbf{z}}_i(\mathbf{x})$[1]. Detailed formulation and analysis of the proposed approach are presented in section II. Numerical simulations are given in section III to validate the proposed approach and the concluding remarks are presented in section IV.

II. Proposed Approach

For simplicity of presentation, the formulation of the proposed approach and the performance analysis in this section are described for a one-dimensional scenario. Furthermore, the nonlinear mappings are considered to be equivalent for each

[1]Please refer to Appendix I for a brief overview for the unscented transformation.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUL 2011		**00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Multi-Sensor Data Fusion: An Unscented Least Squares Approach	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Research Laboratory,Adelphi,MD,20783-1197**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 14th International Conference on Information Fusion held in Chicago, IL on 5-8 July 2011. Sponsored in part by Office of Naval Research and U.S. Army Research Laboratory.

14. ABSTRACT
This manuscript provides an approach to solve the nonlinear least squares problem that arises in decentralized fusion. Even though almost all sensor noise can be modeled as additive noise, the additive nature of the measurement noise is lost when the signal is processed at the sensor node. The proposed approach employs the unscented transformation before the estimation problem at the central node is posed as a nonlinear least squares problem. Numerical simulations indicate that the proposed unscented transformation based approach yields desired results.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

sensor. Overall, the mapping is $\mathbf{h}_i(\cdot) = h(\cdot)$, and as a result, $\hat{\mathbf{z}}_i(\cdot) = \hat{z}(\cdot)$. Consider an observation equation of the form

$$\tilde{y} = f(\tilde{x}) + v \tag{6}$$

It is assumed, the measurement noise is uncorrelated, zero-mean Gaussian noise with known variance, i.e., $v \sim \mathcal{N}(0, \sigma_v^2)$. The sensor node processing can be written as

$$\tilde{z} = h(f(\tilde{x}) + v) + w \tag{7}$$

where $w \sim \mathcal{N}(0, \sigma_w^2)$. Now the problem of interest can be formulated as a nonlinear least squares problem where the cost function to minimize is

$$J(x) = \frac{1}{2N} \sum_{i=1}^{N} \{\tilde{z}_i - \hat{z}(x)\}^2 \tag{8}$$

In the cost function, $\tilde{z}_i$ indicates the individual sensor outputs, i.e.,

$$\tilde{z}_i = h\left(f(\tilde{x}) + v_i\right) + w_i \tag{9}$$

In a traditional approach, a zeroth-order approximation is used to obtain $\hat{z}(x)$ as

$$\hat{z}(x) = h(f(x)) \tag{10}$$

Substituting (10) and (9) into (8) yields the cost function

$$J_1(x) = \frac{1}{2N} \sum_{i=1}^{N} \{h(f(\tilde{x}) + v_i) + w_i - h(f(x))\}^2 \tag{11}$$

Let $y = f(x)$, and $\Delta y = f(\tilde{x}) - f(x)$. Now the above cost function may be written as

$$\begin{aligned} J_1(x) &= \frac{1}{2N} \sum_{i=1}^{N} \{h(y + \Delta y + v_i) + w_i - h(y)\}^2 \\ &= \frac{1}{2N} \sum_{i=1}^{N} \Bigg\{ h(y) + h'(y)(\Delta y + v_i) + \frac{h''(y)}{2!}(\Delta y + v_i)^2 + \\ &\qquad \frac{h'''(y)}{3!}(\Delta y + v_i)^3 + \ldots - h(y) + w_i \Bigg\}^2 \end{aligned}$$

That is, in the traditional approach, the cost function becomes

$$J_1(x) = \frac{1}{2N} \sum_{i=1}^{N} \left\{ w_i + h'(y)(\Delta y + v_i) + \frac{h''(y)}{2!}(\Delta y + v_i)^2 + \frac{h'''(y)}{3!}(\Delta y + v_i)^3 + \ldots \right\}^2 \tag{12}$$

The unscented transformation based approach proposed here uses the sigmapoints[2] to approximate $\hat{z}(x)$, i.e.,

$$\hat{z}(x) = \frac{2}{3}h(f(x)) + \frac{1}{6}h(f(x) + \sqrt{3}\sigma_v) + \frac{1}{6}h(f(x) - \sqrt{3}\sigma_v) \tag{13}$$

[2]Please refer to Appendix I.

Note that here we do not consider sigma points for w since it is assumed to be a zero mean Gaussian variable which appears linearly in $\tilde{z}$. Substituting (13) and (9) into (8) yields

$$\begin{aligned} J_2(x) = \frac{1}{2N} \sum_{i=1}^{N} \Bigg\{ & h(f(\tilde{x}) + v_i) + w_i - \frac{2}{3}h(f(x)) - \\ & \frac{1}{6}h(f(x) + \sqrt{3}\sigma_v) - \frac{1}{6}h(f(x) - \sqrt{3}\sigma_v) \Bigg\}^2 \end{aligned} \tag{14}$$

Using the following Taylor series approximations

$$\begin{aligned} h(f(x) + \sqrt{3}\sigma_v) &= h(y) + h'(y)(\sqrt{3}\sigma_v) + \\ &\quad \frac{h''(y)}{2!}(\sqrt{3}\sigma_v)^2 + \frac{h'''(y)}{3!}(\sqrt{3}\sigma_v)^3 + \ldots \\ h(f(x) - \sqrt{3}\sigma_v) &= h(y) - h'(y)(\sqrt{3}\sigma_v) + \\ &\quad \frac{h''(y)}{2!}(\sqrt{3}\sigma_v)^2 - \frac{h'''(y)}{3!}(\sqrt{3}\sigma_v)^3 + \ldots \end{aligned}$$

yields

$$\begin{aligned} J_2(x) = \frac{1}{2N} \sum_{i=1}^{N} \Bigg\{ & h(y) + h'(y)(\Delta y + v_i) + \frac{h''(y)}{2!}(\Delta y + v_i)^2 \\ & + \frac{h'''(y)}{3!}(\Delta y + v_i)^3 + \ldots - \frac{2}{3}h(y) \\ & - \frac{1}{6}\left[h(y) + h'(y)(\sqrt{3}\sigma_v) + \frac{h''(y)}{2!}(\sqrt{3}\sigma_v)^2 \right. \\ & \left. + \frac{h'''(y)}{3!}(\sqrt{3}\sigma_v)^3 + \ldots \right] \\ & - \frac{1}{6}\left[h(y) - h'(y)(\sqrt{3}\sigma_v) + \frac{h''(y)}{2!}(\sqrt{3}\sigma_v)^2 \right. \\ & \left. - \frac{h'''(y)}{3!}(\sqrt{3}\sigma_v)^3 + \ldots \right] + w_i \Bigg\}^2 \end{aligned}$$

Thus, for the proposed approach, the cost function becomes

$$\begin{aligned} J_2(x) = & \frac{1}{2N} \sum_{i=1}^{N} \Bigg\{ w_i + h'(y)(\Delta y + v_i) + \frac{h''(y)}{2!}(\Delta y + v_i)^2 \\ & + \frac{h'''(y)}{3!}(\Delta y + v_i)^3 + \ldots - \frac{1}{3}\left[\frac{h''(y)}{2!}(\sqrt{3}\sigma_v)^2 \right. \\ & \left. + \frac{h^{(4)}(y)}{4!}(\sqrt{3}\sigma_v)^4 + \frac{h^{(6)}(y)}{6!}(\sqrt{3}\sigma_v)^6 + \ldots \right] \Bigg\}^2 \end{aligned} \tag{15}$$

The accuracy of the least squares solution depends on which approximation of $\hat{z}(x)$ is used in the cost function. The material presented in rest of this section aims to argue that the unscented transformation based approximation given in (13) should yield a superior solution when compared to that of the traditional approximation given in (10).

Lemma 1. *For the least squares problem given in* (8)*, the optimal $\hat{z}(x)$ approaches $E[\tilde{z}]$ as $N \to \infty$.*

Proof: The first derivative of (8) is given by

$$\frac{\partial J}{\partial x} = \frac{1}{N} \sum_{i=1}^{N} \{\hat{z}(x) - \tilde{z}_i\} \frac{\partial \hat{z}}{\partial x}$$

and

$$\frac{\partial J}{\partial x}(x) = 0 \quad \Rightarrow \quad \hat{z}(x) = \frac{1}{N}\sum_{i=1}^{N} \tilde{z}_i$$

satisfies the necessary and the sufficient condition for optimality. Now based on the law of large numbers [6], [7],

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \tilde{z}_i = E[\tilde{z}]$$

Thus the optimal $\hat{z}(x)$ asymptotically approaches $E[\tilde{z}]$, i.e.,

$$\lim_{N\to\infty} \hat{z}(x) = E[\tilde{z}]$$

■

There exists no exact analytical representation for evaluating $E[h(\tilde{y})]$ for a generic nonlinear mapping, $h(\cdot)$. Following the same argument given in [8], it can be shown that the unscented transformation approximation given in (13) is more accurate than the first-order approximation given in (10) when the estimation errors are ignored, i.e., $x = \tilde{x}$. When $x = \tilde{x}$, $\Delta y = 0$ and $E[\tilde{z}]$ can be written as

$$E[\tilde{z}] = h(y) + \frac{h''(y)}{2!}\sigma_v^2 + \frac{h^{(4)}(y)}{4!}3\sigma_v^4 + \frac{h^{(6)}(y)}{6!}15\sigma_v^6 + \frac{h^{(8)}(y)}{8!}105\sigma_v^8 + \dots$$

Using the Taylor series expansion, the unscented transformation approximation can be written as

$$\frac{2}{3}h(f(x)) + \frac{1}{6}h(f(x)+\sqrt{3}\sigma_v) + \frac{1}{6}h(f(x)-\sqrt{3}\sigma_v) = h(f(x)) + \frac{h''(y)}{2!}\sigma_v^2 + \frac{h^{(4)}(y)}{4!}3\sigma_v^4 + \frac{h^{(6)}(y)}{6!}3^2\sigma_v^6 + \frac{h^{(8)}(y)}{8!}3^3\sigma_v^8 + \frac{h^{(10)}(y)}{10!}3^4\sigma_v^{10} + \dots$$

Now the difference between the true expectation and the unscented transformation approximation can be written as

$$E[\tilde{z}] - \left\{\frac{2}{3}h(f(x)) + \frac{1}{6}h(f(x)+\sqrt{3}\sigma_v) + \frac{1}{6}h(f(x)-\sqrt{3}\sigma_v)\right\} = \frac{h^{(6)}(y)}{6!}(15-3^2)\sigma_v^6 + \frac{h^{(8)}(y)}{8!}(105-3^3)\sigma_v^8 + \frac{h^{(10)}(y)}{10!}(945-3^4)\sigma_v^{10} + \dots$$

Also, the difference between the true expectation and the first-order approximation can be written as

$$E[\tilde{z}] - h(f(x)) = \frac{h''(y)}{2!}\sigma_v^2 + \frac{h^{(4)}(y)}{4!}3\sigma_v^4 + \frac{h^{(6)}(y)}{6!}15\sigma_v^6 + \frac{h^{(8)}(y)}{8!}105\sigma_v^8 + \dots$$

Note that for the unscented transformation, the errors only show up in sixth and higher-order terms in the Taylor series. Also, the higher-order errors are scaled. If the sixth and higher-order even terms in the Taylor series of $h(\cdot)$ are identically zero, then the unscented transformation approximation is an exact representation of the true expectation. Therefore, the proposed unscented transformation based approach better approximates $E[\tilde{z}]$ for $x = \tilde{x}$. Thus, we expect that the unscented transformation yields better estimates when compared to the traditional first-order approximation based approach.

From lemma 1, as $N \to \infty$, the optimal solution that minimizes the cost function in (14) is obtained when

$$E[\tilde{z}] - \left\{\frac{2}{3}h(f(x)) + \frac{1}{6}h(f(x)+\sqrt{3}\sigma_v) + \frac{1}{6}h(f(x)-\sqrt{3}\sigma_v)\right\}\Big|_{x=\hat{x}_2} = 0 \tag{16}$$

where $\hat{x}_2$ denotes the x that minimizes (14). Note

$$\begin{aligned} E[\tilde{z}] - \frac{2}{3}h(f(\hat{x}_2)) - \frac{1}{6}h(f(\hat{x}_2)+\sqrt{3}\sigma_v) - \frac{1}{6}h(f(\hat{x}_2)-\sqrt{3}\sigma_v) \\ = h'(\hat{y}_2)E\left[(\Delta y_2 + v)\right] + \frac{h''(\hat{y}_2)}{2!}E\left[(\Delta y_2+v)^2 - \sigma_v^2\right] + \\ \frac{h'''(\hat{y}_2)}{3!}E\left[(\Delta y_2+v)^3\right] + \frac{h^{(4)}(\hat{y}_2)}{4!}E\left[(\Delta y_2+v)^4 - 3\sigma_v^4\right] \\ + \frac{h^{(5)}(\hat{y}_2)}{5!}E\left[(\Delta y_2+v)^5\right] + \dots, \end{aligned}$$

where $\hat{y}_2 = f(\hat{x}_2)$ and $\Delta y_2 = f(\tilde{x}) - f(\hat{x}_2)$. Since v is assumed to be zero mean, Δy_2 can be factored from the odd terms in the Taylor series, i.e.,

$$\begin{aligned} h'(\hat{y}_2)E\left[(\Delta y_2+v)\right] &= \Delta y_2 h'(\hat{y}_2) \\ \frac{h'''(\hat{y}_2)}{3!}E\left[(\Delta y_2+v)^3\right] &= \Delta y_2 \frac{h'''(\hat{y}_2)}{3!}\left\{\Delta y_2^2 + 3\sigma_v^2\right\} \\ \frac{h^{(5)}(\hat{y}_2)}{5!}E\left[(\Delta y_2+v)^5\right] &= \Delta y_2 \frac{h^{(5)}(\hat{y}_2)}{5!}\left\{\Delta y_2^4 + 10\Delta y_2^2\sigma_v^2 + 15\sigma_v^4\right\} \\ \frac{h^{(7)}(\hat{y}_2)}{7!}E\left[(\Delta y_2+v)^7\right] &= \Delta y_2 \frac{h^{(7)}(\hat{y}_2)}{7!}\left\{\Delta y_2^6 + 21\Delta y_2^5\sigma_v^2 + 105\Delta y_2^2\sigma_v^4 + 105\sigma_v^6\right\} \\ &\vdots \end{aligned}$$

Also note that Δy_2 can be factored from the first two even terms in the Taylor series, i.e.,

$$\begin{aligned} \frac{h''(\hat{y}_2)}{2!}E\left[(\Delta y_2+v)^2 - \sigma_v^2\right] &= \frac{h''(\hat{y}_2)}{2!}\Delta y_2^2 \\ \frac{h^{(4)}(\hat{y}_2)}{4!}E\left[(\Delta y_2+v)^4 - 3\sigma_v^4\right] &= \frac{h^{(4)}(\hat{y}_2)}{4!}\left\{\Delta y_2^2 + 6\sigma_v^2\right\}\Delta y_2^2 \end{aligned}$$

Therefore, if the sixth and higher-order even terms in the Taylor series are identically zero, then the optimality condition given in (16) yields $\Delta y_2 = 0$. Now if the mapping $f(\cdot)$ is assumed to be one to one [9], then

$$\Delta y_2 = 0 \quad \Rightarrow \quad \hat{x}_2 = \tilde{x}$$

Thus, the proposed approach asymptotically yields unbiased estimates if the sixth and higher-order even terms in the Taylor series of $h(\cdot)$ about $\hat{y}_2$ are identically zero.

In general, the optimality condition given in (16) can be written in summation form as

$$\mathcal{J}_2(\Delta y_2) = \sum_{k=1}^{\infty} \left\{ \frac{h^{(2k-1)}(\hat{y}_2)}{(2k-1)!} E\left[(\Delta y_2 + v)^{2k-1}\right] + \frac{h^{(2k)}(\hat{y}_2)}{(2k)!} E\left[(\Delta y_2 + v)^{(2k)} - \frac{\{\sqrt{3}\sigma_v\}^{2k}}{3}\right] \right\} = 0 \tag{17}$$

Let $\hat{x}_1$ denotes the optimal x that minimizes (11). Now for the optimal solution that minimizes the cost function in (11) is obtained when

$$\mathcal{J}_1(\Delta y_1) = \sum_{k=1}^{\infty} \left\{ \frac{h^{(2k-1)}(\hat{y}_1)}{(2k-1)!} E\left[(\Delta y_1 + v)^{2k-1}\right] + \frac{h^{(2k)}(\hat{y}_1)}{(2k)!} E\left[(\Delta y_1 + v)^{2k}\right] \right\} = 0 \tag{18}$$

where $\hat{y}_1 = f(\hat{x}_1)$ and $\Delta y_1 = f(\tilde{x}) - f(\hat{x}_1)$. Both optimality conditions given in (17) and (18) do not yield an unbiased solution, unless further constraints are placed on even derivatives of $h(\cdot)$, as mentioned earlier, i.e.,

$$\mathcal{J}_1(0) \neq 0 \quad \& \quad \mathcal{J}_2(0) \neq 0.$$

An optimality condition that yields unbiased estimates can be written as

$$\mathcal{J}^*(\Delta y) = \sum_{k=1}^{\infty} \left\{ \frac{h^{(2k-1)}(y)}{(2k-1)!} E\left[(\Delta y + v)^{2k-1}\right] + \frac{h^{(2k)}(y)}{(2k)!} E\left[(\Delta y + v)^{2k} - E[v^{2k}]\right] \right\} \tag{19}$$

Note that $\mathcal{J}^*(0) = 0$, i.e., $\mathcal{J}^*(\Delta y)$ has a root at the origin. Now subtracting $\mathcal{J}^*(\Delta y)$ from $\mathcal{J}_1(\Delta y)$ yields

$$\begin{aligned} \mathcal{J}_1(\Delta y) - \mathcal{J}^*(\Delta y) &= \sum_{k=1}^{\infty} \frac{h^{(2k)}(y)}{(2k)!} E[v^{2k}] \\ &= \sum_{k=1}^{\infty} \frac{h^{(2k)}(y)}{(2k)!} (2k-1)!! \sigma_v^{2k-1} \end{aligned} \tag{20}$$

Also, subtracting $\mathcal{J}^*(\Delta y)$ from $\mathcal{J}_2(\Delta y)$ yields

$$\mathcal{J}_2(\Delta y) - \mathcal{J}^*(\Delta y) = \sum_{k=1}^{\infty} \frac{h^{(2k)}(y)}{(2k)!} \left\{ E[v^{2k}] - \frac{\{\sqrt{3}\sigma_v\}^{2k}}{3} \right\} \tag{21}$$

Since the measurement noise v is zero mean Gaussian, the first two terms of the Taylor series in (21) vanishes so that:

$$\begin{aligned} \mathcal{J}_2(\Delta y) - \mathcal{J}^*(\Delta y) &= \sum_{k=3}^{\infty} \frac{h^{(2k)}(y)}{(2k)!} \left\{ (2k-1)!! - \frac{\{\sqrt{3}\}^{2k}}{3} \right\} \sigma_v^{2k} \end{aligned} \tag{22}$$

Note that equations (20) and (21) can be written as

$$\mathcal{J}_1(\Delta y) - \mathcal{J}^*(\Delta y) = E[\tilde{z}] - h(f(x))$$

$$\mathcal{J}_2(\Delta y) - \mathcal{J}^*(\Delta y) = E[\tilde{z}] - \frac{2}{3}h(f(x)) - \frac{1}{6}h(f(x) + \sqrt{3}\sigma_v) - \frac{1}{6}h(f(x) - \sqrt{3}\sigma_v)$$

All terms in (22) are all smaller than the corresponding terms in (20). Certainly, as the noise variance goes to zero, both $\mathcal{J}_1$ and $\mathcal{J}_2$ are converging to $\mathcal{J}^*$ for all value of Δy. As the noise variance fall below unity, $|\mathcal{J}^*(\Delta y) - \mathcal{J}_2(\Delta y)| \leq |\mathcal{J}^*(\Delta y) - \mathcal{J}_1(\Delta y)|$ because of the first two terms in (20) that vanish in (22) and the magnitude of all terms in (22) are less than those in (20). Furthermore, it is reasonable to expect that for most cases, the root of $\mathcal{J}_2$ occurs closer to the root of $\mathcal{J}^*$, which happens to be closer to zero than the root of $\mathcal{J}_1$. The intuition is that $\mathcal{J}_2$ is simply a closer surrogate to $\mathcal{J}^*$ than $\mathcal{J}_1$. In short, Δy_2 that satisfies the optimality condition (17) is less than the Δy_1 that satisfies the optimality condition in (18), i.e.,

$$|\Delta y_2| \leq |\Delta y_1|$$

Now we have

$$|f(\tilde{x}) - f(\hat{x}_2)| \leq |f(\tilde{x}) - f(\hat{x}_1)| \tag{23}$$

In general, the inequality condition (23) does not always imply $\hat{x}_2$ is more accurate than $\hat{x}_1$, i.e.,

$$|f(\tilde{x}) - f(\hat{x}_2)| \leq |f(\tilde{x}) - f(\hat{x}_1)| \nRightarrow |\tilde{x} - \hat{x}_2| \leq |\tilde{x} - \hat{x}_1| \tag{24}$$

The superiority of the proposed approach can be obtained if one assumes $\hat{x}_1$ and $\hat{x}_2$ are within an interval, $[-\delta, +\delta]$, centered about $\tilde{x}$ and the function, $f(\cdot)$, can be linearly approximated in that interval.

Table I
SUMMARY OF UNSCENTED LEAST SQUARES APPROACH

Signal	$\tilde{\mathbf{z}}_j = \mathbf{h}\left(\mathbf{f}(\tilde{\mathbf{x}}) + \mathbf{v}_j\right)$ $\mathbf{v}_j \sim \mathcal{N}(\mathbf{0}, R), \quad j = 1, \ldots, N$
UT	$\boldsymbol{\mathcal{Y}}_0 = \mathbf{f}(\mathbf{x}), \quad W_0 = \kappa/(m+\kappa)$ $\boldsymbol{\mathcal{Y}}_i = \mathbf{f}(\mathbf{x}) \pm \left(\sqrt{(m+\kappa)R}\right)_i;$ $W_i = \kappa/\{2(m+\kappa)\}$ $\boldsymbol{\mathcal{Z}}_i = \mathbf{h}(\boldsymbol{\mathcal{Y}}_i)$ $\hat{\mathbf{z}}(\mathbf{x}) = \sum_{i=0}^{2m} W_i \boldsymbol{\mathcal{Z}}_i$
LS	$\arg\min_{\mathbf{x}} \quad J(\mathbf{x}) = \frac{1}{2N} \sum_{j=1}^{N} \{\tilde{\mathbf{z}}_j - \hat{\mathbf{z}}(\mathbf{x})\}^T \{\tilde{\mathbf{z}}_i - \hat{\mathbf{z}}(\mathbf{x})\}$

A summary of the proposed approach for an m-dimensional scenario is given in Table I. The first step in unscented transformation involves obtaining the $2m+1$ sigma

points, $\{\boldsymbol{\mathcal{Y}}_0, \boldsymbol{\mathcal{Y}}_1, \ldots, \boldsymbol{\mathcal{Y}}_{2m}\}$, and the corresponding weights, $\{W_0, W_1, \ldots, W_{2m}\}$. The individual sigma points are then transformed through the nonlinear function to obtain the corresponding $\boldsymbol{\mathcal{Z}}_i$. Finally, $\hat{\mathbf{z}}(\mathbf{x})$ is obtained as the weighted sum of the transformed sigma points.

III. Numerical Simulations

Performance of the unscented least squares approach is evaluated through detailed numerical simulations. Three different simulations are presented here.

A. Simulation I

For the first simulation, we consider the following one-dimensional example.

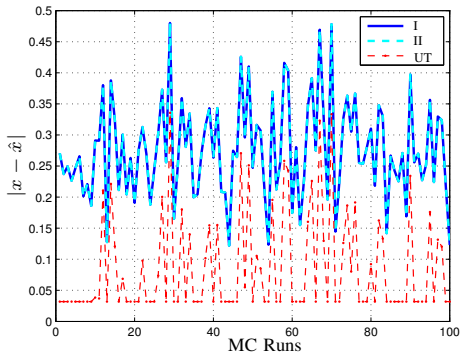
$$\tilde{y}_j = \tilde{x} + v_j, \quad j = 1, \ldots, 10$$
$$\tilde{z}_j = \exp(\tilde{y}_j^2)$$

The measurement noise is selected to be $v_j \sim \mathcal{N}(0, 0.3)$ and the truth is selected to be $\tilde{x} = 0.1/\pi$. The cost function for the least squares is

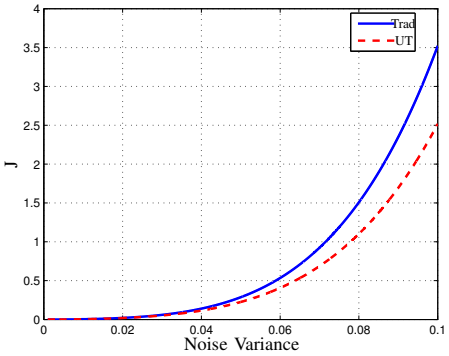
$$J(x) = \frac{1}{2N} \sum_{i=1}^{N} \left\{ \exp((\tilde{x} + v_j)^2) - \hat{z}(x) \right\}^2 \qquad (25)$$

The problem is solved using three different approaches. The first two approaches use the traditional least squares approach, i.e., $\hat{z}(x) = \exp(x^2)$. The third approach uses the proposed unscented transformation based approach. In the first approach, the optimization problem in (25) is solved using the Gauss-Newton method [10]. The second approach uses the Nelder-Simplex algorithm [11], i.e., the *fminsearch* function in Matlab. The third approach also uses the *fminsearch* function to solve the optimization problem obtained from the proposed unscented transformation based scheme.

Results obtained from 100 Monte-Carlo runs are presented in figure 1. The estimation error presented in figure 1(a) indicates that the unscented transformation based approach (UT) yields better results when compared to the traditional approach. Note that the unscented transformation based approach exhibits an error floor of ≈ 0.0318 in figure 1. This error floor is obtained when $\hat{x} = 0$ and it is due to the ambiguity in the nonlinear least squares problem presented. Figure 1(a) also indicates that the iterative least squares approach and the simplex algorithm yield the same estimation error for the traditional approach. The sensitivity of the cost function, or J, to the noise intensity is presented in figure 1(b). Clearly, the traditional cost function is more sensitive to the noise intensity when compared to the proposed approach. Numerical values for the mean and standard deviation of the estimation error as well as the mean and standard deviation of the minimum J are given in Table II. The table indicates that the unscented approach is superior because it produces a less biased estimate. On the other hand, the variance of the error is comparable for the two approaches.



(a) Estimation Error



(b) Noise Sensitivity of Cost

Figure 1. Simulation I: Estimation Error & Noise Sensitivity of Cost Function

Table II
Simulation I: Monte-Carlo Results

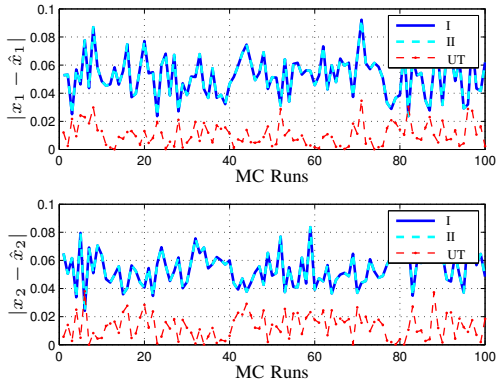
	Traditional Iterative	Traditional *fminsearch*	Proposed *fminsearch*
Mean Error	0.2752	0.2752	0.0835
Error Stdv	0.0796	0.0796	0.0812
Mean Cost	0.3091	0.3091	0.3190
Cost Stdv	0.5509	0.5509	0.5462

B. Simulation II

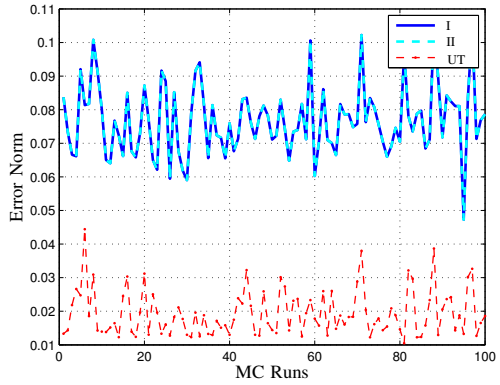
For the second simulation, we consider the following two-dimensional example.

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix}_j = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix}_j + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}_j, \quad j = 1, \ldots, 20$$
$$\begin{bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \end{bmatrix}_j = \begin{bmatrix} \sqrt{\tilde{y}_1^2 + \tilde{y}_2^2} \\ \exp(\tilde{y}_1) + \exp(\tilde{y}_2) \end{bmatrix}_j$$

The measurement noise is selected to be $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \begin{bmatrix} 0.007 & 0 \\ 0 & 0.1 \end{bmatrix})$ and the truth is $\mathbf{x} = \begin{bmatrix} 0.023 \\ 0.038 \end{bmatrix}$.



(a) Estimation Error



(b) Error Norm

Figure 2. Simulation II: Estimation Error & Error Norm

For the second simulation, the results obtained from the Monte-Carlo runs presented in figure 2 indicate that the unscented transformation based approach yields a better estimate when compared to the traditional approach. Figure 2 also indicates that the iterative least squares approach and the simplex algorithm yield a similar estimation error for the traditional approach. Again, the UT exhibits superior estimation performance due to smaller bias.

C. Simulation III

The third simulation considers the practical problem of sniper localization using multiple gunfire detection systems [12], [13]. The shooter or the target location and the i^{th} sensor location are defined as T and S_i, respectively. For simplicity, the problem is formulated in $\Re^2$, i.e., $T \in \Re^2 \equiv \begin{bmatrix} T_x \\ T_y \end{bmatrix}$ and $S_i \in \Re^2 \equiv \begin{bmatrix} S_{i_x} \\ S_{i_y} \end{bmatrix}$. Now we define the individual range, r_i, and bearing, ϕ_i, between the i^{th} sensor node and the target as

$$r_i = \sqrt{(T_x - S_{i_x})^2 + (T_y - S_{i_y})^2} \tag{26}$$

$$\phi_i = \arctan 2 \left(T_y - S_{i_y}, T_x - S_{i_x}\right) \tag{27}$$

Figure 3 illustrates the geometry of the shockwave and the

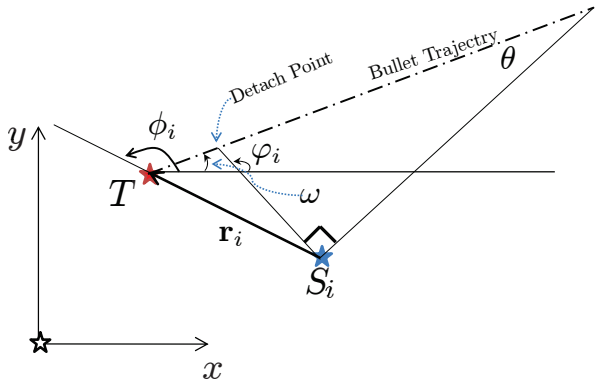


Figure 3. Schematic Representation

muzzle blast for the i^{th} sensor node when the orientation of the bullet trajectory is ω with respect to the horizontal axis. Here the angle ϕ_i indicate the direction of arrival (DOA) of the muzzle blast, φ_i indicate the DOA of the shockwave, and the shockwave cone angle θ is defined as

$$\theta = \arcsin\left(\frac{1}{m}\right) \tag{28}$$

where m is the Mach number. The Mach number is assumed to be known since the typical value for Mach number is $m = 2$. For more detailed description of the scenario, please refer to [12]. Now the DOA angle for the shockwave can be written as

$$\varphi_i = \begin{cases} -\frac{\pi}{2} - \theta + \omega, & \text{if } \pi + \omega < \phi_i < -\frac{\pi}{2} - \theta + \omega; \\ \frac{\pi}{2} + \theta + \omega, & \text{if } \frac{\pi}{2} + \theta + \omega < \phi_i < \pi + \omega. \end{cases} \tag{29}$$

The first case $\pi + \omega < \phi_i < -\frac{\pi}{2} - \theta + \omega$ corresponds to the scenario where the sensor is located above the bullet trajectory and the third case $\frac{\pi}{2} + \theta + \omega < \phi_i < \pi + \omega$ corresponds to the scenario where the sensor is located below the bullet trajectory. The second case $\phi_i = \pi + \omega$ corresponds to the scenario when the sensor is located on the bullet trajectory, and such a scenario is not considered here.

Under the assumptions that the bullet maintains a constant velocity over its trajectory, the time difference between the shockwave and the muzzle blast can be written as [14], [15]

$$\tau_i = \frac{r_i}{c}\left[1 - \cos\left|\phi_i - \varphi_i\right|\right] \tag{30}$$

When the sensor node is within the field of view (FOV) of the shockwave, the three available measurements are the two DOA angles and the time difference of arrival (TDOA) between the muzzle blast and the shockwave, i.e.,

$$\hat{\phi}_i = h_1\left(T, S_i, \omega\right) + \eta_\phi \tag{31}$$

$$\hat{\varphi}_i = h_2\left(T, S_i, \omega\right) + \eta_\varphi \tag{32}$$

$$\hat{\tau}_i = h_3\left(T, S_i, \omega\right) + \eta_\tau \tag{33}$$

where $h_1(\cdot)$ is given in (27), $h_2(\cdot)$ is given in (29) and $h_3(\cdot)$ is given in (30). The measurement noise is assumed

to be zero mean Gaussian white noise, i.e., $\eta_\phi \sim \mathcal{N}(0, \sigma_\phi^2)$, $\eta_\varphi \sim \mathcal{N}(0, \sigma_\varphi^2)$, and $\eta_\tau \sim \mathcal{N}(0, \sigma_\tau^2)$. After obtaining the measurements at each sensors, the sensor range and bearing to the target it calculated as follows:

$$\hat{\phi}_i = \hat{\phi}_i \tag{34}$$

$$\hat{r}_i = \frac{c\hat{\tau}_i}{\left[1 - \cos\left|\hat{\phi}_i - \hat{\varphi}_i\right|\right]} \tag{35}$$

Individual bearing and range information is then relayed to the central node where it is fused along with the sensor locations to yield an accurate sensor location. These information sources are related to the unknown parameters T and ω by a nonlinear transformation that changes for each sensor due the changes in the sensor location (see (31)-(33)). In fact for range, the angular measurement errors falls inside the nonlinear function in (35). For numerical simulation presented here, we consider the scenario presented in figure 4. Measurement noise models

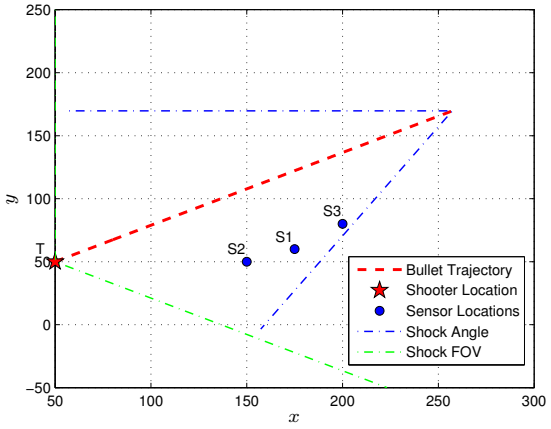
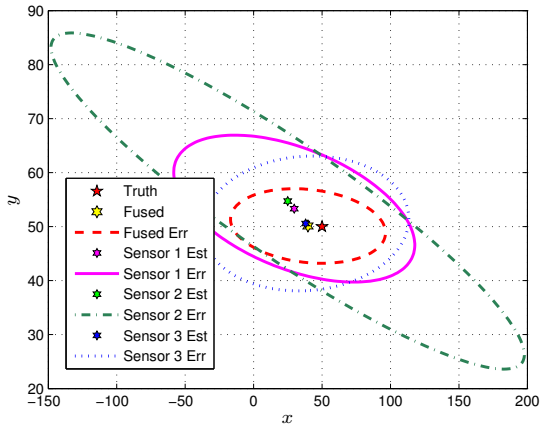


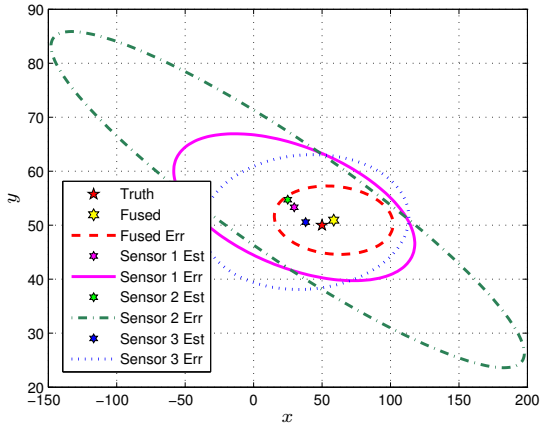
Figure 4. Simulation Scenario

are selected as $\sigma_\phi = \sigma_\varphi = 4^o$ and $\sigma_\tau = 1$msec.

Figures 5(a) and (b) show the individual sensor estimates of the shooter location as well as the fused estimates and the corresponding error ellipses for both the traditional and the unscented transformation approaches. Note that for both approaches, the fusion algorithm provides an estimate that is superior to the individual sensor estimates. Also, note that the fused estimate obtained from the proposed approach is closer to the truth. Numerical values for the true shooter location as well as the estimated locations obtained from both approaches and the corresponding root-mean-square-error (RMSE) are given in Table III. For this scenario, the UT again provides an estimate with smaller bias than the traditional approach. In fact, the error ellipse for the UT approach is smaller than that of the traditional approach. Unlike the prior scenarios, the performance improvement from the UT can be attributed to both reduction in the bias and standard deviation of the error.



(a) Traditional Least Squares



(b) Unscented Least Squares

Figure 5. Simulation III: Estimation Error Ellipse from Monte-Carlo Runs

Table III
SIMULATION III: MONTE-CARLO RESULTS

	x	y
Truth	50	50
Trad. Fused	39.7429	50.1127
UT. Fused	58.5026	50.9283
Trad. RMSE	49.1189	5.8610
UT. RMSE	37.7813	5.4535

IV. CONCLUSION

This manuscript provides an unscented approach to the nonlinear least squares problem that arises in decentralized fusion. In decentralized fusion, measurements are first processed at the sensor node before they are relayed to the central node. Due to the signal processing conducted at the sensor node,

one might lose the additive nature of the measurement noise. The proposed unscented transformation based approach helps to tackle the non-additive nature of the noise in the nonlinear least squares problem. Utilizing the information regarding the noise characteristic, the proposed approach yields superior estimates when compared to results obtained from the traditional least squares approach. Simulation results presented here indicate that unitizing the unscented transformation helps to alleviate the cost sensitivity to the measurement noise. Also, the unscented transformation helps to move the minimum cost corresponding to the optimal solution toward the cost associated with the zero estimation error and thus improves the estimates.

ACKNOWLEDGEMENT

The first author wishes to acknowledge Yang Cheng, Assistant Professor, Department of Aerospace Engineering, Mississippi State University, for the numerous helpful discussions and comments.

APPENDIX I

The unscented transformation is a method for approximating, in a computationally efficient manner, the statistics of a random variable that undergoes a nonlinear transformation. The basic principle behind unscented transformation is that it is easier to approximate a probability distribution than an arbitrary nonlinear function [8]. Consider an n-dimensional random variable, $\mathbf{x}$, with mean $\bar{\mathbf{x}}$ and covariance P_{x}. A second random variable $\mathbf{y}$ is related to $\mathbf{x}$ through the nonlinear function

$$\mathbf{y} = \mathbf{h}(\mathbf{x})$$

To calculate the statistics of the $\mathbf{y}$ using the unscented transform, we proceed as follows: First, a set of $2n+1$ weighted samples or *sigma points* $\mathcal{S}_i = \{W_i, \boldsymbol{\mathcal{X}}_i\}$ are deterministically chosen so that they completely capture the true mean and covariance of the prior random variable $\mathbf{x}$. A selection scheme that satisfies this requirement is

$$\begin{aligned}
\boldsymbol{\mathcal{X}}_0 &= \bar{\mathbf{x}}, \qquad W_0 = \kappa/(n+\kappa) \\
\boldsymbol{\mathcal{X}}_i &= \bar{\mathbf{x}} + \left(\sqrt{(n+\kappa)P_{\mathrm{x}}}\right)_i, \quad W_i = \kappa/\{2(n+\kappa)\}, \\
&\qquad\qquad\qquad\qquad i = 1, \ldots, n \qquad (36) \\
\boldsymbol{\mathcal{X}}_i &= \bar{\mathbf{x}} - \left(\sqrt{(n+\kappa)P_{\mathrm{x}}}\right)_i, \quad W_i = \kappa/\{2(n+\kappa)\}, \\
&\qquad\qquad\qquad\qquad i = n+1, \ldots, 2n
\end{aligned}$$

where κ is a scaling parameter and $\left(\sqrt{(n+\kappa)P_{\mathrm{x}}}\right)_i$ is the i^{th} row or column of the matrix square root of $(n+\kappa)P_{\mathrm{x}}$. The weights are selected so that

$$\sum_{i=0}^{2n} W_i = 1$$

If $\mathbf{x}$ is Gaussian, then selecting κ such that $(n+\kappa) = 3$ minimizes the mean-squared-error up to the fourth order [16]. However, when κ is negative, then there exists a possibility that the predicted covariance can become a non-positive semi-definite. If this is of concern, then another approach can be used that allows for scaling of the sigma points, which guarantees a positive semi-definite covariance matrix [17]. Sigma point are then propagated through the nonlinear function

$$\boldsymbol{\mathcal{Y}}_i = \mathbf{h}(\boldsymbol{\mathcal{X}}_i), \qquad i = 0, \ldots, 2n \qquad (37)$$

and the estimated mean and covariance of $\mathbf{y}$ are computed as follows

$$\bar{\mathbf{y}} = \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{Y}}_i \qquad (38)$$

$$P_{\mathrm{y}} = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}\right)\left(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}\right)^T \qquad (39)$$

These estimates are accurate to the second (third order for Gaussian priors) of the Taylor series expansion of $\mathbf{h}(\cdot)$ for any nonlinear function. Errors introduced in the higher-order terms are scaled by the choice of the parameter κ.

REFERENCES

[1] H. Mitchell, *Multi-Sensor Data Fusion: An Introduction*. New York, NY: Springer-Verlag, 2007.

[2] O. Kessler and F. White, "Data fusion perspectives and its role in information processing," in *Handbook of Multisensor Data Fusion: Theory and Practice*, M. E. Liggins, D. L. Hall, and J. Llinas, Eds. Boca Raton, FL: CRC Press, 2009, ch. 2.

[3] S. J. Julier and J. K. Uhlmann, "A consistent, debiased method for converting between polar and cartesian coordinate systems," in *The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, 1997, pp. 110 – 121.

[4] ——, "A new extension of the Kalman filter to nonlinear systems," in *Proc. SPIE*, vol. 3068, 1997, pp. 182–193.

[5] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401 – 422, March 2004.

[6] J. Jacod and P. Protter, *Probability Essentials*, 2nd ed. Berlin, Germany: Springer-Verlag, 2004.

[7] H. G. Tucker, *A Graduate Course in Probability*. New York, NY: Academic Press, 1967.

[8] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *Automatic Control, IEEE Transactions on*, vol. 45, no. 3, pp. 477 – 482, March 2000.

[9] A. Brown and C. M. Pearcy, *An Introduction to Analysis*. New York, NY: Springer-Verlag, 1995.

[10] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic System*. Boca Raton, FL: Chapman & Hall/CRC, 2004, ch. 1.

[11] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.

[12] L. Kaplan, T. Damarla, and T. Pham, "QoI for passive acoustic gunfire localization," in *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, Oct. 2008, pp. 754 –759.

[13] D. Lindgren, O. Wilsson, F. Gustafsson, and H. Habberstad, "Shooter localization in wireless sensor networks," in *Information Fusion, 2009. FUSION '09. 12th International Conference on*, 2009, pp. 404 –411.

[14] J. Bedard and S. Pare, "Ferret: a small arms fire detection system: localization concepts," vol. 5071, no. 1. SPIE, 2003, pp. 497–509.

[15] T. Damarla, L. Kaplan, and G. Whipps, "Sniper localization using acoustic asynchronous sensors," *Sensors Journal, IEEE*, vol. 10, no. 9, pp. 1469 –1478, 2010.

[16] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *American Control Conference, 1995. Proceedings of the*, vol. 3, 1995, pp. 1628 – 1632.

[17] S. Julier, "The scaled unscented transformation," in *American Control Conference, 2002. Proceedings of the*, vol. 6, 2002, pp. 4555 – 4559.